

Research Summary: Stress Counter-Response Training via Physiological Self-Regulation During Flight Simulation
July 29, 1998 – June 30, 2000**By Olafur S. Palsson, Psy.D., Principal Investigator****Abstract**

This study provided the first evaluation of a new training concept and technology aimed at training pilots to maintain physiological equilibrium during circumstances in an airplane cockpit. Thirty healthy subjects (16 males and 14 females) between the ages of 18 and 35 were randomized into two study groups, A and B. Subjects participated individually in a sequence of four study sessions. In the first visit, subjects were taught to operate a desktop fighter jet flight simulation program. In the three sessions that followed, subjects in group A were trained to minimize their autonomic deviation from baseline values while operating the desktop flight simulation. This was done by making their skin conductance and hand temperature deviations from baseline impair the functionality of the aircraft controls. Subjects also received auditory and visual cues about their autonomic deviation, and were instructed to keep these within pre-set limits to retain full control of the aircraft. Subjects in group B were subjected to periods of impaired aircraft functionality independent of their physiologic activity, and thus served as a control group. No statistically significant group differences were found in the flight performance scores from the three training sessions, and post-training flight performance scores of the two groups were not different. We conclude that this study did not provide clear support for this training methodology in optimizing pilot performance. However, a number of shortcomings in the current status of this training methodology may account for the lack of demonstrable training benefit to the experimental group. Suggested future modifications for research on this training methodology include: Limiting the amount of instrument impairment resulting from physiological deviations; conducting a greater number of physiological training sessions per subject; using pre-post training performance tests which invoke a greater amount of stress in subjects; and developing a more detailed performance scoring system.

Introduction

The majority of accidents in modern aviation can be traced to human error. As sophistication in automated aircraft control systems, with multiple backups and automated emergency responses, has steadily increased and has made the machine aspects of flying more reliable, the reliability of the human aviator has increasingly become a limiting factor in efforts to further increase aviation safety. The proposed study, described below, will serve as a preliminary study to examine the effectiveness of a novel training concept for reducing pilot error during demanding or unexpected events in the cockpit by teaching pilots self-regulation of excessive autonomic nervous system (ANS) reactivity during simulated flight tasks.

Egorov (1982) has defined the aviator's professional reliability as the ability to handle flight task demands satisfactorily in limited time and solve any problems in emergency conditions. Zhang et al (1997) suggest that this professional reliability depends on two relative factors: (1) task demand load, and (2) the pilot's cognitive functional capacity. Piloting in modern aviation is a complex problem-solving task composed of many subtasks which compete for limited cognitive processing capacity. Although routine aircraft operation may keep these two factors sufficiently balanced for reliable pilot operation, the balance can shift if stressful flight events occur. During unexpected events or emergencies in the cockpit, the demand for fast-paced and non-routine mental problem solving (or task demand load) increases while the perceived threat of the situation and/or cognitive strain produce an emotional response which produce excessive autonomic arousal which may reduce the pilot's capacity for effective cognitive performance (functional capacity).

The relationship between sympathetic ANS activation and cognitive performance has been recognized for nearly a century, and is most popularly expressed in the Yerkes-Dodson law

(Yerkes & Dodson, 1908) which postulates that performance of cognitive tasks is best under a moderate level of stress (this is often depicted graphically as an inverted U where performance is the Y axis and emotional stress or sympathetic arousal as the X axis). Research has shown that the optimal levels of stress for effective performance vary with task complexity and familiarity with the task. Zhang et al. (1997) have found that the total cognitive processing capacity of professional pilots is inversely correlated with emotional anxiety during flight. These investigators used the Psychophysiological Reserve Capacity (PRC), a measure of the pilot's ability to accomplish tasks in addition to the primary flight tasks, to evaluate the total cognitive processing capacity of the subjects.

It is well established in psychophysiological research that there are vast differences in the degree to which individuals respond to the same stressor with increase in sympathetic ANS activity. These interpersonal differences have been found to be partly dependent on personality factors which determine a person's interpretation of life events. For example, a recent study in our laboratory has shown the personality variables of Extraversion, Absorption and Social Desirability significantly affect the strength of cardiovascular responses to a standard cognitive stressor (Palsson et al, 1998). In aviation, individual pilots vary substantially in their autonomic reactivity to demanding flight tasks (Comens, Reed, & Mette, 1987). Personal cognitive traits may contribute greatly to such individual differences in the stress reactivity of pilots (as well as their judgment) by effecting over assessment or underassessment of the consequences of non-routine events (Simmel, Cerkovnik & McCarthy, 1989).

Physical stress is a collective term for changes in the activity of numerous physiological processes which occur in the presence of, or in response to, perceived threat or challenge. This response (the so-called fight-or-flight response), which is aimed at preparing the organism for effective physical response to the threat, is triggered in the hypothalamus and carried to various organ systems by the sympathetic part of the autonomic nervous system (neural axis) as well as through a hormonal chain reaction (the hypothalamic-pituitary-adrenergic axis). The changes in cognitive performance in response to stress are the result of several aspects of the physiological stress response, including changes in brain blood flow and changes in blood sugar levels. However, the most immediate and most potent modulating factor is direct ascending neural activation of higher brain centers through the reticular formation. The reticular formation is a complex network of around 100 nuclei that run from the top of the brain stem to the bottom of the myelencephalon. Various nuclei in the reticular formation are involved in cognitive arousal, but also in many other basic functions, such as movement, muscle tone maintenance, cardiac reflexes, circulation, and modulation of attention.

The amount of activation in the neural part of the sympathetic stress response can be measured relatively reliably through surface measurement of hand temperature and hand skin conductance. Both the arterioles (the small blood vessels which determine the amount of blood flow to the skin) and the skin sweat glands receive neural input exclusively from the sympathetic part of the autonomic nervous system. Decrease in hand temperature and increase in sweating therefore generally reflect increased sympathetic arousal level. Both of these measures have been widely used in clinical biofeedback for several decades, and will be used for training subjects to limit their autonomic reactivity during simulated flight in the proposed study.

Specific Aims:

This study evaluated a new training concept for training pilots for maintaining physiological equilibrium suited for optimal cognitive and motor performance under emergency events in an airplane cockpit. The training method is novel in that it (a) adapts biofeedback methodology to train physiological balance during simulated operation of an airplane, and (b) uses graded impairment of control over the flight task to encourage the pilot to gain mastery over his/her autonomic functions. This training concept has been termed Instrument Functionality Feedback (IFF) and was developed jointly between NASA (Alan T. Pope, Ph.D.) and EVMS (Olafur S. Palsson, Psy.D.), and was tested and refined in a preliminary way in this research study.

Subjects:

Thirty healthy subjects between ages of 18 and 35 (14 females and 16 males, mean age = 25.6 years) completed four individual experiment sessions in the EVMS Behavioral Medicine Clinic. All recruited subjects were free from anxiety disorders or other stress-related health problems, and without history of attention difficulties. Subjects were randomly divided into two groups of 15 individuals. Subjects were recruited until 30 subjects had completed participation in all four study sessions. Any subjects who did not complete participation in all four sessions, or did not show sufficient ability to learn the flight simulation task in the first visit, were replaced through additional recruitment (a total of 7 subjects were replaced in this manner).

Subjects were compensated \$20 per completed session and could earn an additional bonus for each of the three training sessions depending on their flight performance success. Additionally, all parking costs at EVMS was paid for subjects.

Preparations for data collection:

The first twelve months of the project consisted of preparation for data collection. The main accomplishments during this phase were as follows:

1. Choosing flight simulation software which was suitable for this kind of research. Numerous off-the-shelf flight software programs were tested before Jane's Anthology of Fighters was selected as the best suited for training, based on the requirements described above;
2. Programming and testing the flight scenarios to be used in the study;
3. Solving various problems with the protocol. This included adding sound and lights for biofeedback in addition to the instrument functionality modulation, and changing the protocol from norm-referenced physiologic values as basis for deviation calculations, to individualized normalization in the beginning of each session;
4. Developing a standardized flight performance scoring system;
5. Recruitment of subjects; and
6. Developing a consistent flight instruction method to teach subjects to fly in the first training session.

Study methods:**Equipment and software:**

Subjects trained in all four sessions with an off-the-shelf flight simulation software (Jane's Anthology of Fighters) which ran on an IBM compatible computer. Selection of the software used for this study was based on the following criteria: 1. adequate mastery of the game can be accomplished by the average naïve operator within one hour; 2. an accurate scoring system which provides reliable measure of the pilot's performance could be developed; and 3. the software could provide adequate challenge to place significant stress load on the operator.

The technology under investigation was a prototype of a customized biofeedback input system developed at NASA LaRC, which consisted of a joystick, Colbourn physiological modules and a J & J I-330 biofeedback software running on a second IBM compatible computer. The subjects operated the simulated aircraft via joystick with one hand, and temperature sensor (placed on the palmar surface of the little finger) and skin conductance sensors (metal plates on middle and ring fingers) were attached to the other (non-dominant) hand. Subjects also used a keyboard to control various aspects of the aircraft's operation from time to time. The functionality of the joystick action upon the game (side-to-side or forward/backward movement as well as the use of the trigger) was modulated by the biofeedback input system in a pre-set way adjustable by the trainer, and in the experimental group (A) this functionality was inversely related to deviations in hand temperature and skin conductance (see below) measured by the sensors on the other (free) hand of the subject.

The modulation of joystick functionality was different for the two study groups:

Group A (the experimental training group) had temperature and skin conductance levels baseline values manually set by the trainer at take-off, based on the trainees present functioning, and the physiologically influenced joystick modulation equipment impaired aircraft controls in the following manner throughout the three training sessions:

- a. finger temperature deviations beyond 1 degree Fahrenheit away from session baseline resulted in graded impairment, culminating in total loss of roll and pitch joystick capability when a deviation of 3 degrees Fahrenheit was reached. One of two warning lights indicating the direction of temperature deviation beyond 1 degree F (blue for low temperature or red for high temperature) alerted the trainee to temperature deviation.
- b. skin conductance deviations greater than 1 micromho away from session baseline resulted in weapons (the F-16's cannon and missiles) to be off-line until the deviation was corrected. This was indicated to the trainee through a central light which lit up during such deviation.

Group B (the control group) was subjected to 2-minute reduction in functionality at pre-set times. Hence, these impairments were independent of pilot's physiological state and were introduced into the task of this control group to equalize frustration levels and flight challenge levels between the groups due to impaired cockpit controls. During Flight Segment I in sessions 2-4, joystick functionality was reduced by 25% for two minutes after reaching the first way-point, and 50% after reaching the second way-point. In Flight Segment II of each of these sessions, where fighting took place, trainees in group B were subjected to 1 minute of 50% reduction in pitch and roll ability and loss of weapons after each 5 minutes of normal joystick functionality.

Experiment Protocol

Subjects completed 4 flight sessions in four different visits, all of which were generally completed within a period of 2-4 weeks. Subjects flew simulated F-16 aircraft operating in a desert daytime environment. The flight missions both included simple take-off, navigation, and landing, and fighting mission where the trainee was alone in the sky against three equally capable enemy aircraft. The sequence of sessions and flight tasks for each subject was as follows:

Practice Session (Visit 1). The goal of the practice session is to give the subjects adequate opportunity to learn full control of the flight task sufficient for comfortable normal operation of the simulated aircraft. The practice session consisted of 75 minutes of structured guided instruction in the operation of the aircraft simulation, both in normal operation (take-off, navigation of a pre-set course, and landing) and in fighting enemy aircrafts. During this initial guided instruction phase, subjects were instructed in a standardized manner in the operation and controls of the aircraft simulation, and practiced flying with correction from the instructor when needed.

Pre-Training Flight Test (Visit 1):

After a short break following the Practice Session described above, subjects underwent a Pre-training Flight Test. During this test, they completed a take-off, flight to two way-points at specified altitudes and landing within a pre-programmed flight scenario, and were given a performance score on a standardized rating scale. Subjects who had not gained full comfortable control of the flight task at the end of the practice period, as evidenced by a low score (lower than 7 out of 13 points) on this flight test, were eliminated from the study.

Training Sessions (Visits 2-4):

Each of the three training sessions will consist of two segments:

Segment I: The flight tasks of this 5-10 minute flight round were correct take-off, successful navigation to two pre-programmed destinations, and correct return and landing.

Segment II: The flight tasks of this segment consisted of up to 45-minutes of continuous flight with changes in temperature and skin conductance causing variations in degree of joystick

functionality for group A, but pre-determined intervals of instrument impairment for group B, as described above. During each Segment II period, trainees entered into air combat with three enemy planes, and their performance was scored partly based on their combat capability.

Post-Training Flight Test (Visit 4)

After a brief break following the third training session, a post-training flight test was given to each trainee. This test was identical to the Pre-training Flight Test.

Standardized scoring sheets were used by the trainer to rate the trainee's performance in each session. The performance of the trainees was manually documented on these sheets by the trainer, according to pre-set criteria.

Data analysis and results:

The hypotheses guiding data analysis and interpretation of results were tested using independent-group t-tests. Three hypotheses were tested:

Hypothesis I: Performance scores on a standardized demanding flight task will be better for group A than for group B. This was assessed by comparing the performance on a 13 point rating scale of flight performance on the post-training flight test between the two groups. The hypothesis was not supported in the two-group t-test. This comparison is shown in Figure 1. Post training performance scores for groups A (experimental training group) and B (control group) were not found to be statistically significantly different.

Hypothesis II: Precision in flight performance will be greater after training in Group A than in Group B (due to better physiological self-regulation). This was evaluated by comparing how close to the exact target altitudes the trainees were when they reached each of the first two way-points in the post-training flight test; the trainees were instructed to try to keep their aircraft as close as they possibly could to 4000 feet and 2000 feet upon reaching the first and second way-point, respectively. As seen in Figures 2 and 3, no group differences were seen in the precision of this altitude task. Hypothesis II was therefore not supported by the data.

Hypothesis III: Flight performance will improve steadily across the three flight training sessions in both groups, but the rate of improvement will be greater from session to session for Group A. As demonstrated by Figure 4, this hypothesis was not statistically supported in the data analysis. No group differences were seen in performance in any of the three flight training sessions. However, there seems to be a visible statistically non-significant trend in the data for a steeper improvement in flight performance for Group A than Group B. In Figure 4, this can be seen from the fact that group A has a lower mean value in session II, but higher mean value for session IV. Larger group size or a greater number of training sessions might have made this trend statistically significant.

Conclusions and future research directions:

We conclude that the present study failed to provide any clear evidence for an effect of the tested methodology or technology on optimizing pilot performance under challenging circumstances. However, this lack of significant training effect may have more to do with the particular implementation of the "Instrument Functionality Feedback (IFF)" methodology in this first test of the IFF concept than with the training concept itself. Several possible problems were identified in the course of this research, and further studies should aim to rectify these potential shortcomings or confounds in their implementation of IFF methodology.

1. *Unlimited instrument functionality impairment:* In this study, no limits were placed on how much the aircraft controls could be impaired by physiological deviation from baseline. Hence, temperature deviation of three degrees from baseline resulted in total loss of control of the aircraft roll and pitch, which generally resulted in a crash. Such total loss of control was stressful to the subjects and may have induced a sense of helplessness that was not conducive to self-regulation efforts.
2. *The low number of training sessions.* Subjects in Group A only had three sessions to learn physiological self-regulation as well as practicing a novel and complex flight task. It

seems quite likely in retrospect, that this was an inadequate amount of practice time to see the effect of the physiological training. In traditional biofeedback, it often takes at least three sessions to begin to master self-regulation, and it may take longer when such self-regulation demand is coupled with the mastery of another complex task. A longer sequence of training sessions, perhaps six or eight session, might have yielded significant positive training results. That notion seems to be supported by the trend (which did not reach statistical significance) for Group A to improve more than Group B across the three training sessions (see Figure 4).

3. *Using a pre-post training performance which invokes little stress in subjects.* In this study, the post-training flight test simply consisted of take-off, navigation, and landing. This nature of the test was chosen to allow full standardization of the task (battle or emergency events are harder to standardize). It is likely that this was not an adequately stressful task to call upon the physiologic learning (or lack thereof) in a way that made a difference evident in the flight performance). Although difficult to implement in a standardized way, it is recommended that future studies utilize highly challenging or stressful events in the post-training flight test.
4. *Using a more refined flight performance scoring method.* The scoring method used in the present study to evaluate post-training flight performance, rated the trainees' performance on only 13 different points of optimal flying characteristics. The majority of trainees scored in the range from 11 and 13 on the post-training flight test, and this scoring system therefore turned out to be a weak measure of individual differences. A more comprehensive scoring system with greater resolution for assessing individual differences and individual progress should be developed for future studies.

Figure 1. Mean Post-training flight performance scores for the experimental training group (A) and control group (B). Group performance was not significantly different.

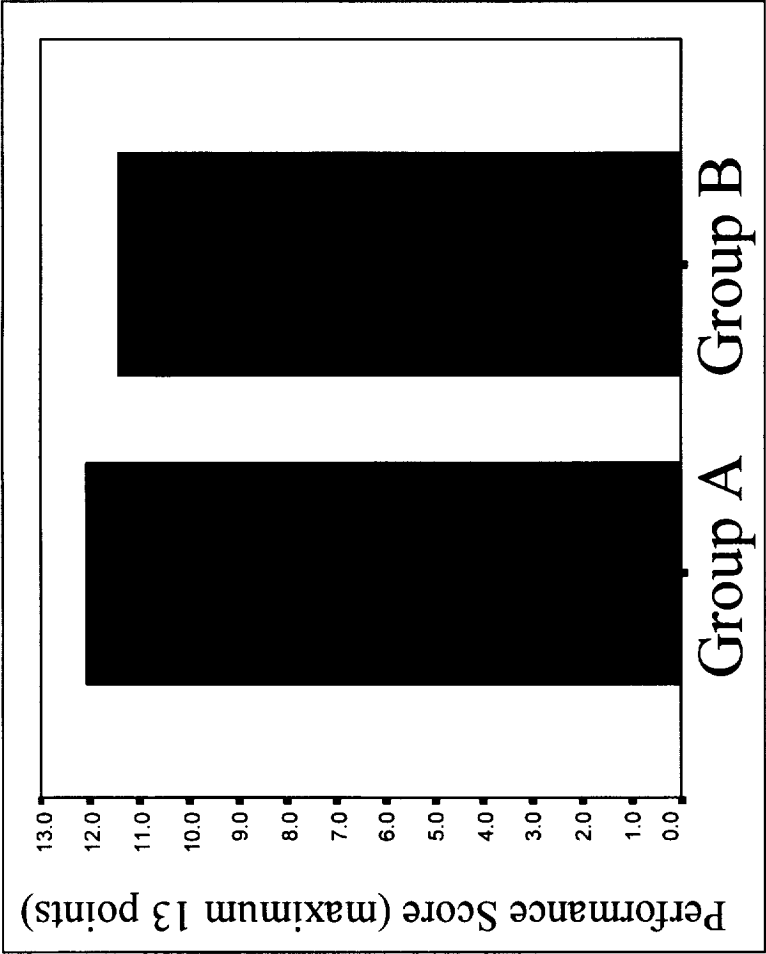


Figure 2. Mean deviation in feet from target altitude upon reaching the first way-point on the post-training flight test. The groups did not show statistically significant difference in this measure of flight precision.

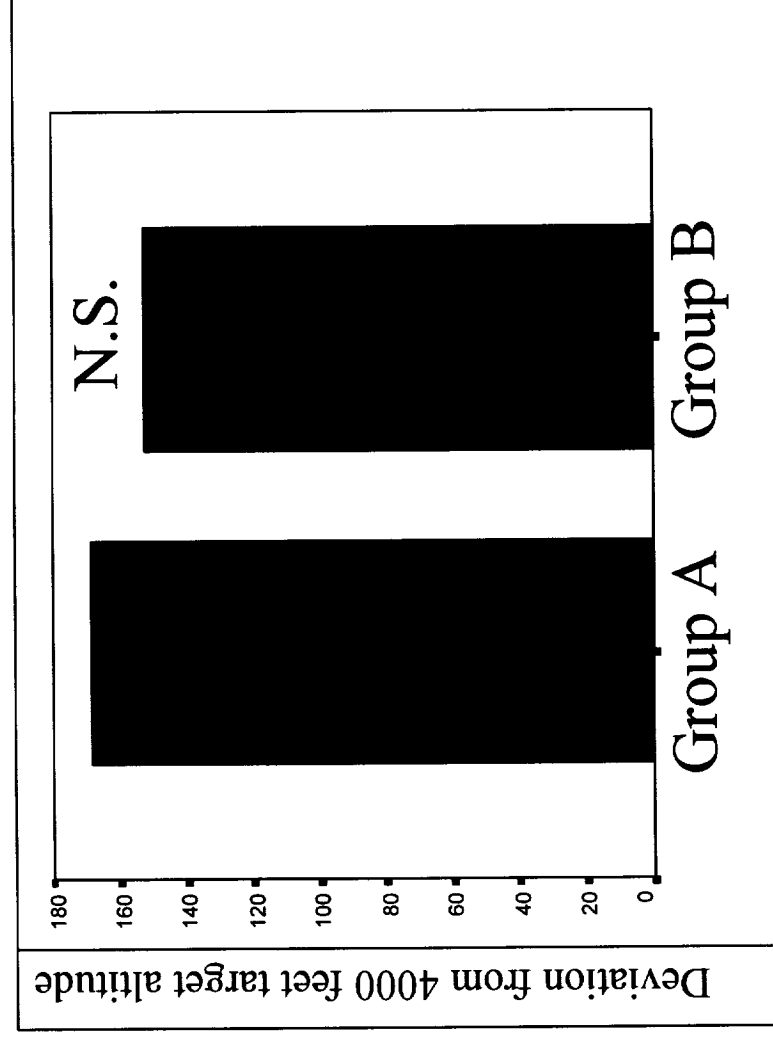


Figure 3. Mean deviation in feet from target altitude upon reaching the second waypoint on the post-training flight test. The groups did not show statistically significant difference in this measure of flight precision.

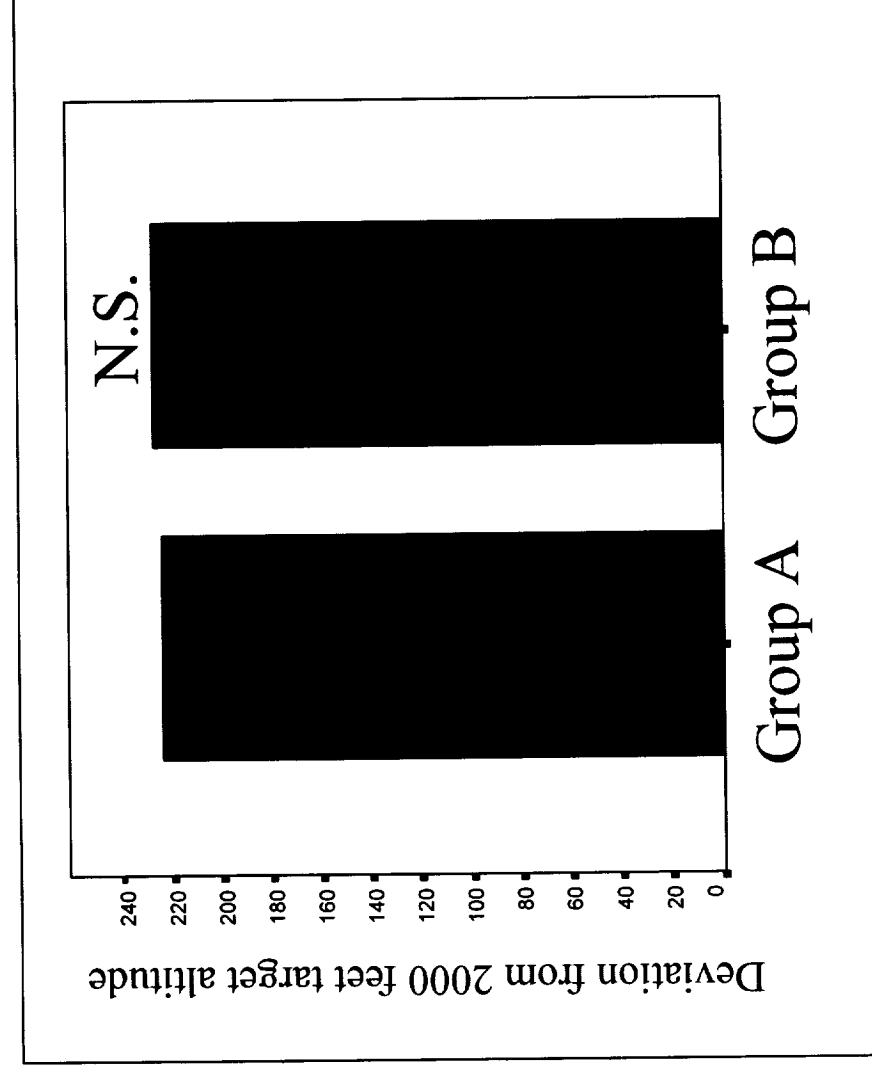
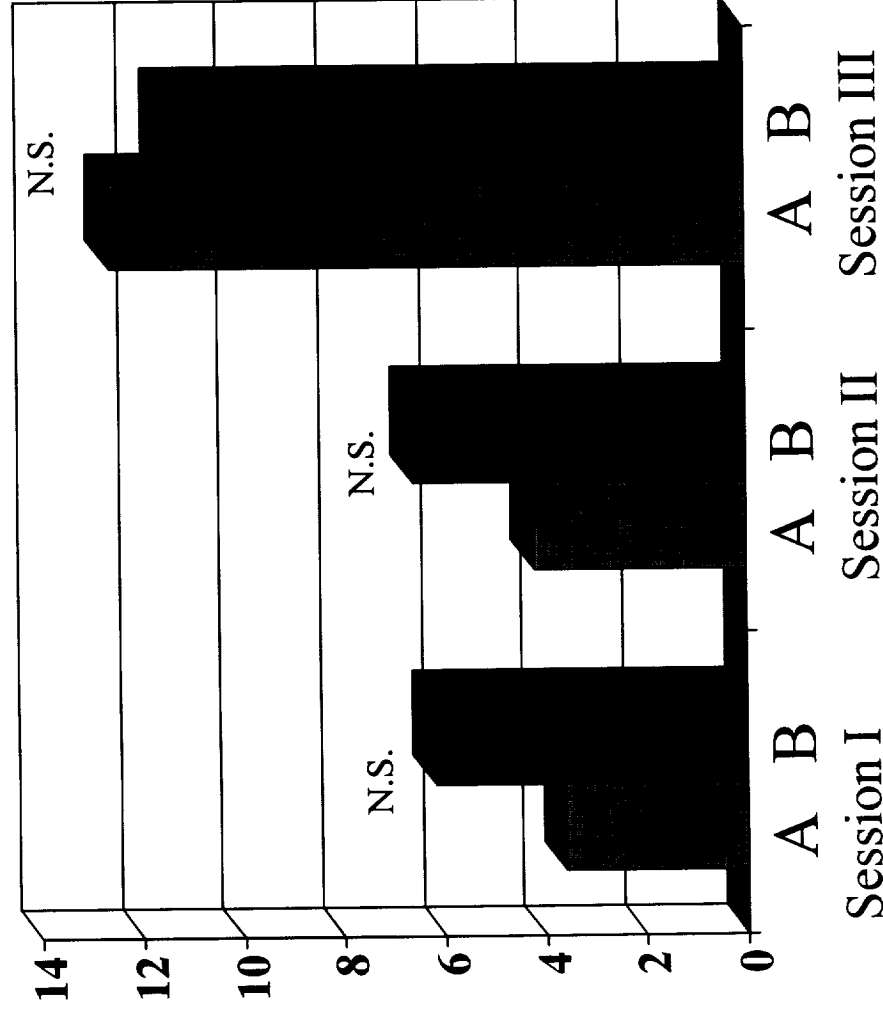


Figure 4. Mean flight session performance in the three training sessions for groups A and B. No group differences were statistically significant.



Groups and Sessions